VCRIX - volatility index for crypto-currencies on the basis of CRIX

Master Thesis Submitted to

Prof. Dr. Wolfgang K. Härdle
Prof. Dr. Cathy Yi-Hsuan Chen

School of Business and Economics
Ladislaus von Bortkiewicz Chair of Statistics

Humboldt-Universität zu Berlin

by

Alisa Kolesnikova

585968

in partial fulfillment of the requirements for the degree of

Master of Economics and Management Science

Berlin, May 30, 2018

Acknowledgement

I would like to express my sincere gratitude to my supervisor Prof. Dr. Wolfgang K. Härdle for support, understanding and inspiration to do my best, that he has been providing not only during the production of this thesis, but throughout my whole master degree.

I would also like to thank Prof. Dr. Cathy Yi-Hsuan Chen for her advice and great opportunities to learn from her during my time at the LvB Chair.

My sincere gratitude extends to Dr. Simon Trimborn who has provided extensive support and guidance, his advice played a great role in the production of this thesis.

Additionally I would like to thank Prof. PhD. Hermann Elendner and Dr. Andrea Barletta for offering their expertise and providing extensive feedback, their valuable input helped to shape this research. I wish to extend my thanks to all of those who supported me in any respect during the completion of the thesis, in particular all members of the Ladislaus von Bortkiewicz Chair of Statistics.

Alisa Kolesnikova

Abstract

The crypto-currency market brings along unusual levels of risk and returns compared to traditional markets. Historical volatility shows behavior that is rendering established trading strategies obsolete calling for new methods of portfolio management, yet the implied volatility is of even higher importance since it allows for an expectation about the future risk. For the US market CBOE offers the VIX, in Germany VDAX is provided as a measure for implied volatility, both based on the respective options market. Given the absence of a developed crypto-currency derivatives market, this research proposes a methodology to create VCRIX (a parent index to CRIX) - a volatility index, able to grasp the risk induced by the crypto-currency market. VCRIX addresses the market dynamics providing a mean directional accuracy (MDA) of 46% compared to CRIX realized volatility estimated from high-frequency data. In an application of the methodology to the US market, we are able to track the performance of VIX with correlation of 69 % and an MDA of 60%. VCRIX is shown to be an adequate measure for implied volatility, thus proved to be a proper basis for option pricing. The codes used to obtain the results in this paper are available via www.quantlet.de

Keywords: index construction, model selection, volatility, crypto-currency, VCRIX

JEL classification: C51, C52, G10

Contents

List of Figures

List of Tables

1 Introduction

Since the inception of Bitcoin in 2008 the crypto ecosystem has seen a market capitalization explosion that reached 795 billion USD at its highest point on the 6th of January 2018(coinmarketcap.com). Following the adoption of crypto-currencies, several governments announced plans to introduce state cryptos, more than 2.3 billion USD was raised with initial coin offerings according to icowatchlist (2018)(for comparison, notorious Snapchat IPO on March 1, 2017 raised around 3.4 billion USD), and even widely known skeptics like JP Morgan launched the Blockchain Center of Excellence (BCOE) for exploration of blockchain potential in finance industry. Regardless of a major market meltdown in the beginning of 2018, crypto-currencies remain on the watch list of governments, industry leaders and academics.

Apart from traditional hedge-funds and institutional investors who are interested in diversification, the crypto ecosystem saw more than 200 crypto-funds incepted during the last 3 years (Autonomous Research Report). Growing interest led to investigation efforts on the behavior of this new asset class. The classification of crypto-currencies remains a topic of heated debate: currencies, assets, commodities or investment vehicles (Yermack (2015), Glaser et al. (2014)). A large share of the existing research treats crypto-currencies in terms of financial assets traded much like common assets for investment and speculative purposes. This paper follows the same paradigm.

Those not believing in the blockchain technology underlying most of cryptos, were enticed by outstanding returns and risk management opportunities, which led to the development of new portfolio strategies (Elendner et al. (2016)) and attempts to further explore the potential of crypto-currencies as an investment tool. However, traditional market instruments (indexes, ratings, investment portfolios) were not ready to accommodate for the new asset class, leaving the crypto-currencies out of scope for an average investor. The CRyptocurrency IndeX developed by Trimborn and Härdle (2016) was designed to grasp the dynamics of the young crypto-currency market using liquidity-based rules and information criterion. Components of the index are revised every 3 months. The number of revised constituents changed greatly over time reflecting the tectonic changes of the crypto ecosystem taking place over 2017-2018, these changes can be seen in Figure1.1:

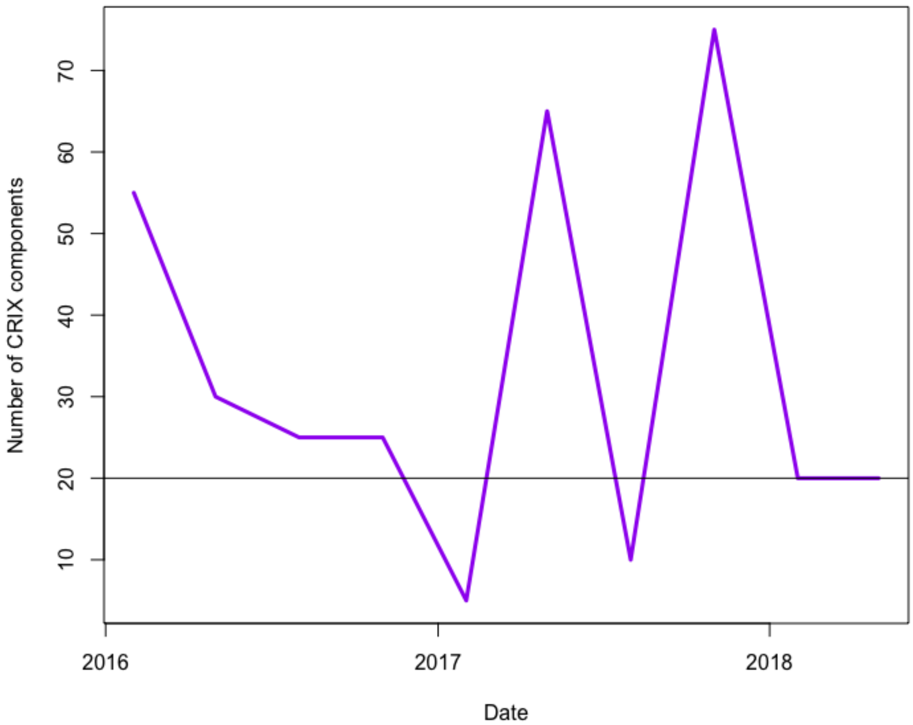

Figure 1.1: Number of CRIX components over time.

The early months of 2016 are driven by the introduction of a larger number of coins and general quietness of the market which allowed newcomers to have their say. The lower number of constituents signalizes that market movements are driven by the jumps of BTC, ETH and XRP (like in the "awakening" phase of Feb of 2017 or political turmoil of Aug 2017). Higher numbers of constituents normally signify the increasing role of the altcoins in the global exchange. Here and further on in this paper "crypto-currencies" and "cryptos" should be considered equal terms.

1.1 Motivation

Rapid growth of Bitcoin price led to persistent talks about 'bubble-like' behavior and general skepticism of the market (Baek and Elbeck (2015), Cheung et al. (2015)), exposing the need for deeper understanding of the underlying processes driving the valuation of crypto-currencies. Research in this field was done by Hayes (2017) and White (2015). Introduction of Bitcoin futures by the CME and CBOE on December 18, 2017 reinforced the positions of crypto-currencies as a new asset class. The emergence of the derivative market signalled the need for the solid pricing strategies as well as a reliable risk measure. This need was addressed in the paper on pricing crypto-currencies by Chen et al. (2018), where the Stochastic Volatility with Correlated Jumps model (based on methodology proposed by Duffie et al. (2000) and implied volatility dynamics study by Fengler et al. (2003)) is offered as a response to previously discovered non-stationarity and local inhomogeneity of CRIX returns. The paper provides a framework for option pricing and revealed the necessity to further explore the behavior of the CRIX volatility, so as to provide the final ingredient - a proxy for an implied volatility. This paper will make aims to fulfill this gap. The goal of the proposed index (VCRIX) is the estimation of the risk measurement for the CRIX components and delivery of market status information to potential investors.

Before the detailed explanation of the index methodology is provided in chapter 2, the paper will offer a brief reminder on the history of implied volatility estimation efforts and an overview of the existing traditional volatility indexes in section 1.3. Apart from the index design, the VCRIX methodology chapter will include a revision of CRIX structure (section 2.1) and its specifics as well as an overview of the motivation behind the choice of variance estimation model (section 2.3). VCRIX remains a backward-looking index, as it will be constructed using historical volatility, however, the employed methodology can be potentially used in order to make valid forecast on the market volatility. Apart from the general time series analysis of the estimated index, section 2.4 will provide a deeper exploration of the decay parameter intuition (section 2.5) and explore the connection between VCRIX behavior and development of the crypto ecosystem. Chapter 3 will offer insights into the VCRIX performance with regard to two benchmarks: high-frequency estimation and VIX simulation. The concluding chapter 4 will summarize the findings of the paper and look into potential development paths for VCRIX.

1.2 Implied volatility

Implied volatility became a subject of academic research with the development of the derivative market, mostly due to the complications of its estimation. With the revolutionary option pricing model introduced by Black and Scholes (1976), the market crash of October 1987, that bent the volatility surface of index options into a skewed "volatility smile" (it became a varying function of strike and expiration, according to Derman and Kani (1994)), rendered Black-Scholes (BS) model invalid, and after 40 years there is still no overwhelming consensus on the "correct" model. BS model suggested that implied volatility should be the best predictor of future volatility because, by definition, the implied volatility is the future volatility expected by the market. Future volatility is a crucial component for asset pricing and risk management. It matters to option traders, insurance companies, and any company whose business model relies on risk analysis. Every market had to settle on the model most appropriate for the specifics of the corresponding derivatives behavior. It is only reasonable that crypto-currency market will have to find its own model that will reflect the instability that it is currently experiencing. Consideration of the existing volatility indexes would constitute a logical step towards the selection of the appropriate solution.

1.3 Volatility indexes

Expected future volatility plays a major role in finance theory and can be crucial to adequate decision making. Volatility indexes are key measures of market expectations of volatility implicit in the prices of options (generally, option implied volatility is estimated by the inverse function of the BS formula). The so-called "fear indexes" convey the investor sentiment and offer a valuable insight into the market expectations.

Model-free indexes seem to be dominating the field. A recent summary of implied volatility indexes provided by Siriopoulos and Fassas (2009) considers official market indexes of countries that have highly liquid options markets and readily available model-free implied variances (France, Germany, Japan, Switzerland, the U.K., and the U.S). In this research, the VIX by CBOE in USA would be considered the closest counterpart of the VCRIX. The current VIX methodology was developed on the basis of pioneering research of Neubürger (1994), Madan et al. (1998), Demeterfi et al. (1999) and Britten-Jones and Neuberger (2000) among others. It is showcasing the implied volatility calculated from the option prices on derivative market based on S&P 500 by taking strike prices and option prices as inputs. More particularly, VIX became proxy for market volatility, using market-implied variance swap rate, calculated based on the exchange-traded S&P 500 variance swaps (CBOE (2009)):

$$\sigma^2 = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T}\left[\frac{F}{K_i} - 1\right]^2 \tag{1.1}$$

Where $\sigma = VIX/100$, T is time to expiration, F is a forward index level from index option prices, K_o is a first strike price below F and K_i is a strike price of the ith out-of-money option.

The model-free implied volatility (MFIV) constitutes the basis of VIX and can be hard to measure with accuracy due to the lack of precise prices for options with strikes in the tails of the return distribution. Andersen and Bondarenko (2007) provide a derivation of the corridor implied volatility measure and relate it to MFIV under general assumptions, offering an alternative estimation model for the implied volatility.

In case of crypto-currencies there is no developed derivative market yet. Bitcoin futures are being traded on CBOE, Ethereum derivatives might follow, but for now, there is no options data for the rest of the market. The nascent stage of the crypto derivative market creates the major complexity for the construction of the volatility index that would be capable of reflecting more than just a standard deviation of major cryptos. As an extrapolation of the existing methodology was rendered impossible, a new methodology for the best approximation to the expected volatility of CRIX basket had to be developed, given the specifics of the crypto-currency time series behavior.

2 VCRIX

The current empirical research will develop a methodology that will offer the estimation of the implied volatility given the absence of the derivative market. One of the first dichotomies to consider was the choice between model-free or model-driven approach. Recent years saw the rise of the model-free indexes (based on "model-free" implied volatility (MFIV)), that could arguably stem from a certain disappointment in BS model. Major contestants are the Black-Sholes (BS) implied volatility and statistical models such as GARCH. MFIV is extracted from the corresponding set of current option prices without the need to assume any specific pricing model. This approach is certainly attractive however, it comes along with a range of methodological issues in calculating and using MFIV. Biktimirov and Wang (2017) tested both approaches on the subject of forecasting accuracy, and surprisingly BS implied volatility came out superior both in terms of in-sample "encompassing" models that include several forecasts in the same combined specification and also in out-of-sample forecasting. Following the mentioned limitations on data availability (no derivatives on cryptos available, Bitcoin futures data would be non-informative for estimation of the entire market) and the expected performance of model-based indexes, VCRIX has to be based on a model, capable of processing the particularities of the underlying assets: liquidity, strong correlation among the assets, highly inhomogeneous returns history due to adoption cycles, non-stable range of constituents. Fulfillment of these requirements and additional rationale behind them will be further explored in this section.

2.1 Components and structure of CRIX

S&P500 and DAX provide a summary statistic of the current state of their respective markets. CRIX became a similar indicator for the crypto-currency market, providing a statistically-backed market measure, unlike other crypto indexes like Crypto20, CCi30, WorldCoinIndex. For this reason, it was selected as the best basis for the volatility index of cryptos that this paper proposes.

Returns on the crypto-currencies that constitute CRIX at corresponding periods will provide input for VCRIX, thus the index rules will have a significant impact on the behaviour of VCRIX. The original paper Trimborn and Härdle (2016) defines CRIX as a Laspeyres index, taking the value of a k asset basket and comparing it against the base period. The final formula constitutes an adjusted formula of Laspeyres:

$$CRIX_t(k,\beta) = \frac{\sum_{i=1}^{k} \beta_{i,t_l^-} P_{it} Q_{i,t_l^-}}{Divisor(k)_{t_{l^-}}} \tag{2.1}$$

with P_{it} the price of asset i at time t and Q_{i,t_l^-} the quantity of asset i at time t_l^-(the last time point when Q_{i,t_l^-} was updated), β_{i,t_l^-} the adjustment factor of asset i at time point t_l^- with l^{th} being the adjustment factor. For market indexes, such as CRSP, SP500 or DAX, the quantity Q_{i0} is the number of shares of the asset i in the base period. The *Divisor* ensures that the monthly re-balancing that accounts for the changes in a market volume of crypto-currencies and the number of index components, affects the value of the CRIX in a consistent manner and prevents structural jumps. At the starting point of the CRIX, the *Divisor* is chosen such that:

$$Divisor_1 = \frac{\sum MV_{i0}}{1000} \tag{2.2}$$

Traditional approach dictates a constant number of constituents. However, the crypto market is under influence of many non-conventional factors like technological innovation, political uncertainty and level of adoption among others. Additionally, due to almost zero costs of creating a token, new participants appear almost daily. In order to gain best perspective on the market dynamics with less variance, CRIX uses a selection procedure based on the Akaike Information Criterion and Schwartz

Information Criterion to choose the crypto-currencies that are the most representative of the crypto ecosystem.

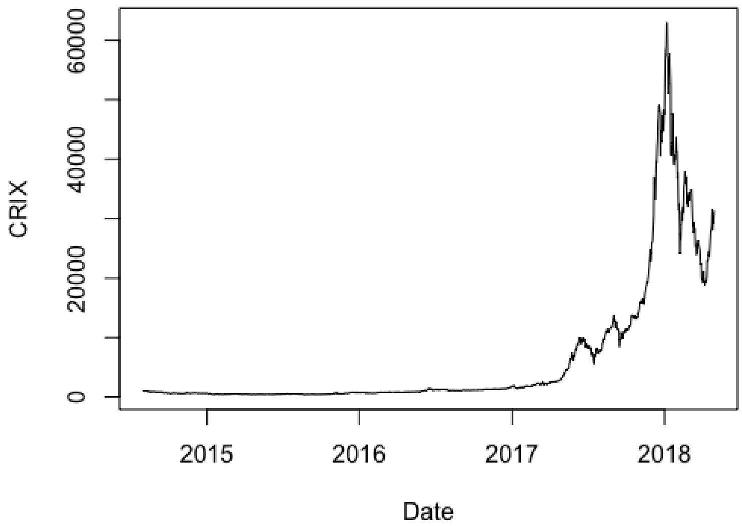

Figure 2.1: CRIX.

CRIXcode

In order to account for new coins and volatility of existing constituents, index components are revised quarterly. As previously shown in Table 1.1, the number of constituents can vary greatly from quarter to quarter, this change will be reflected in the calculation of VCRIX as well.

The daily historic values of CRIX constituents for the period of May 2016 to May 2018 were sourced from crix.hu-berlin.de and converted to returns as shown in equation 2.4:

$$R_{t+1} = \frac{price_{t+1}}{price_t} \tag{2.3}$$

$$r_{t+1} = ln(R_{t+1}) = ln(price_{t+1}) - ln(price_t) \tag{2.4}$$

The returns on CRIX constituents were further used for variance estimation as described in subsection 2.3

2.2 Index structure

One of the major purposes of volatility estimation is the risk management - to measure the potential losses, estimates must be made of future volatilities and correlations. The development of the blockchain ecosystem and integration of crypto-currencies led to the development of complex relationships between the cryptos:

	BCH	BTC	DASH	EOS	ETC	ETH	IOTA	LTC	XMR	XRP
BCH	1									
BTC	0.24	1								
DASH	0.24	0.36	1							
EOS	0.32	0.41	0.22	1						
ETC	0.09	0.12	0.10	0.12	1					
ETH	0.25	0.45	0.30	0.37	0.19	1				
IOTA	0.21	0.41	0.25	0.41	0.12	0.35	1			
LTC	0.26	0.54	0.29	0.35	0.14	0.42	0.37	1		
XMR	0.24	0.46	0.34	0.28	0.12	0.42	0.36	0.40	1	
XRP	0.17	0.25	0.14	0.31	0.06	0.25	0.22	0.33	0.25	1

Table 2.1: Correlation of returns on top 10 crypto-currencies.

As depicted in the table above, most of the currencies have a positive correlation with Bitcoin, Ethereum and Monero. These currencies remain in the spotlight of the media for a long time and could be considered drivers of the crypto-currency adoption. Thus, relationships between the CRIX constituents could potentially have a significant impact on the volatility of the index.

The Markowitz (1952) approach of minimizing risk for a given level of expected returns has become a standard approach and was employed in VCRIX in order to account for the effect of correlation between returns on the volatility of CRIX basket and reflect the weights of the constituents derived in CRIX estimation.

The volatility index is designed to grasp the dynamics of returns on crypto-currencies as well as their internal relationships by accounting for the covariances of the constituents:

$$VCRIX = \frac{\sqrt{\sum_i^n w_i \sigma_i^2 + \sum_i \sum_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij}}}{Divisor} \tag{2.5}$$

Components *i,j* of the volatility index are defined by returns on the constituents of CRIX. The weights w reflect market capitalization of the respective crypto-currencies and match those of CRIX (reset every 3 months). σ^2 signifies **daily** variance, while ρ is defining the correlation between crypto returns, thus making $\sigma_i \sigma_j \rho_{ij}$ a covariance between returns on crypto-currency i and crypto-currency j on the selected day t. The estimation method for the σ will be covered in detail in the following subsection.

The initial value of VCRIX will be set to 1000, following the convention set by CRIX, as described in subsection 2.1. A *Divisor* is introduced in order to account for the jumps that might occur due to the change in the number of constituents every month. The *Divisor* is set to a certain value on the first day to transform the estimated volatility to 1000 points of VCRIX. *Divisor* remains the same over the month. In the formula below the *basket* defines the set of crypto-currencies that constituted the index at day t:

$$1000 = VCRIX_1 = \frac{\sigma_{basket,1}}{Divisor_1} \tag{2.6}$$

$$\frac{\sigma_{basket,t-1}}{Divisor_{t-1}} = VCRIX_{t-1} = VCRIX_t = \frac{\sigma_{basket,t}}{Divisor_t} \tag{2.7}$$

Every month the number of constituents changes, in this case, the value of VCRIX from the last day of the month will be transferred to the first day of the next month, after that the *Divisor* will be reevaluated in order to reflect the value for transformation of the first day volatility to the value of VCRIX transferred from the last day of previous month. The new *Divisor* will be used until the end of the next month. So far the *Divisor* has been taking values from $4.1 \cdot 10^{-5}$ to $9.1 \cdot 10^{-5}$.

2.3 Variance-covariance estimation

The selection of the variance-covariance matrix estimation method for the returns of the index components is by design standing at the center of the methodology. A number of attempts to find the model bringing us closest to the true volatility was undertaken in the past. Modeling efforts have been made by Catania et al. (2018), favoring the performance of the IGARCH and GJGARCH models. A paper on the first econometric analysis of CRIX family time series by Chen et al. (2016) suggests the ARIMA (2,0,2)-t-GARCH(1,1) model with normally distributed residuals. The paper registered the volatility clustering phenomenon when looking at CRIX returns and heavy tail features. Most of current research favors the GARCH-based models for the volatility prediction, due to the observed specifics of the time series: volatility autocorrelation, heteroscedasticity of returns, fat tails, asymmetric reactions on upward and downward movements. However, global models have a disadvantage that this paper accounted for: the dynamics of crypto-currency market, its inhomogeneous history with long years of adoption dictates a need for a local model that would discount the results of the earlier phase and assign larger weights to the recent movements. VCRIX is using historical volatility as an input, however, its goal is to come as close to implied volatility as possible and gain certain predictive power. Research conducted on forecasting ability of ARCH models out-of-sample in non-USA markets (Dimson and Marsh (1990), Tse (1991), Kuen and Hoong (1992)), argues that simpler models like exponentially weighted moving average (EWMA) performed better in volatility forecasting tasks. EWMA methods have been around since the 1950s, and are still the most popular forecasting methods used in business and industry in the forecasting literature according to Hyndman and Shang (2009). The difference between the GARCH model developed by Engle (1982)and Bollerslev (1986) and the EWMA model is analogous to the difference between equation 2.8 and 2.9:

$$\sigma_t^2 = \sum_{i=1}^{m} \alpha_i r_{t-i}^2 \tag{2.8}$$

$$\sigma_t^2 = \lambda\omega + \sum_{i=1}^{m} \alpha_i r_{t-i}^2 \tag{2.9}$$

EWMA is a particular case of GARCH, that comprises elimination of long-run average variance rate (by setting its weight to 0). Taking into account the stochastic nature of volatility and extreme cases that are observed when analyzing the crypto-currencies, the EWMA was considered a more logical method of variance estimation, especially given the potential use of VCRIX for forecasting. The standard EWMA assumes a normal distribution of returns. In practice returns on crypto-currencies are often skewed and heavy-tailed like returns of traditional financial assets (Hull and White (1998)). Therefore Lu et al. (2010) suggestes a robust-EWMA based on a Laplace distribution. Robust-EWMA offers a certain adjustment, where volatility $\sigma_{i,t+1}$ is an EWMA version of the maximum likelihood estimator of the standard deviation, $\hat{\sigma} = \frac{1}{n}\sum_{t=1}^{n}\sqrt{2}|r_t|$ in the traditional Laplace distribution. However, this method will be put aside for further research, meanwhile, the variance estimation through standard EWMA is defined as follows:

$$\sigma_{i,t+1}^2 = \lambda\sigma_{i,t}^2 + (1-\lambda)r_{i,t}^2 \tag{2.10}$$

$$\sigma_{ij,t+1} = \lambda\sigma_{ij,t} + (1-\lambda)r_{i,t}r_{j,t} \tag{2.11}$$

$$\rho_t = \frac{\sigma_{ij,t}}{\sigma_{i,t}\sigma_{j,t}} \tag{2.12}$$

where $\sigma_{i,t+1}^2$ is the variance of returns (r) of crypto-currency i in the next period, σ_{ij} is a covariance of cryptos i and j and ρ_t is a correlation between two cryptos in time t ($0 < \lambda < 1$ remains a decay factor, it will be explored in more detail in subsection 2.5).

In general EWMA method allows the estimation of the local rather than the global level of the variance, following the assumption that in the crypto market. The influence of the earliest observations should have almost no effect on the most recent ones, thus allowing the estimation of variances and covariances based on the method similar to a rolling window, with application to the decay factor λ. Additionally, JP Morgan suggests EWMA for forecasting the conditional volatility of short-horizon returns in terms of conditional variance, that could be potentially useful for the use of VCRIX for forecasting.

Before moving to the evaluation of the method and benchmark tests, the next section will explore the main characteristics and the stylized facts of the VCRIX, as well as link the events that took place during the observation period to the behavior of time series on the graph.

2.4 Time series analysis of VCRIX

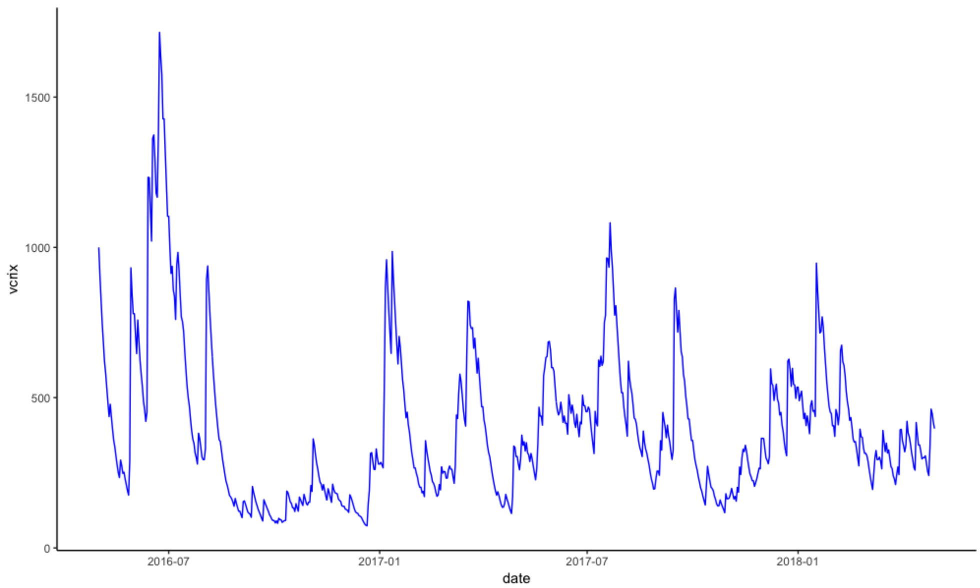

Figure 2.2: VCRIX May 2016 - May 2018.

VCRIXindex

As Figure 2.2 shows, from May 2016 to May 2018 daily VCRIX displayed a mean value of 398 points and a standard deviation of 257 index points. The maximum level of 1715 points was reached by the end of June 2017, with the minimum being 59 points by the end of 2016. Additionally, an interesting trend could be observed in Figure 2.3: following the initial spike and drop, the smoothing pattern was growing during 2017, however, it seems to be stabilizing by the last quarter of 2017 at 380 points (approximately 5.4% daily volatility). Crypto markets are known to break the usual patterns, nonetheless, the observed stabilization might signal the drift to maturity of the ecosystem. Fitting of the VCRIX time series suggests the ARIMA (1,0,1) process with non-zero mean. Augmented Dickey-Fuller test rejects the null hypothesis in favor of stationarity at 1% significance level. VIX index also displays stationarity, however, as it will be mentioned in subsection 2.5, local inhomogeneity is an expected feature of modelled volatility. This analysis is out of the scope of this thesis, however, it certainly constitutes necessary steps for the research to follow.

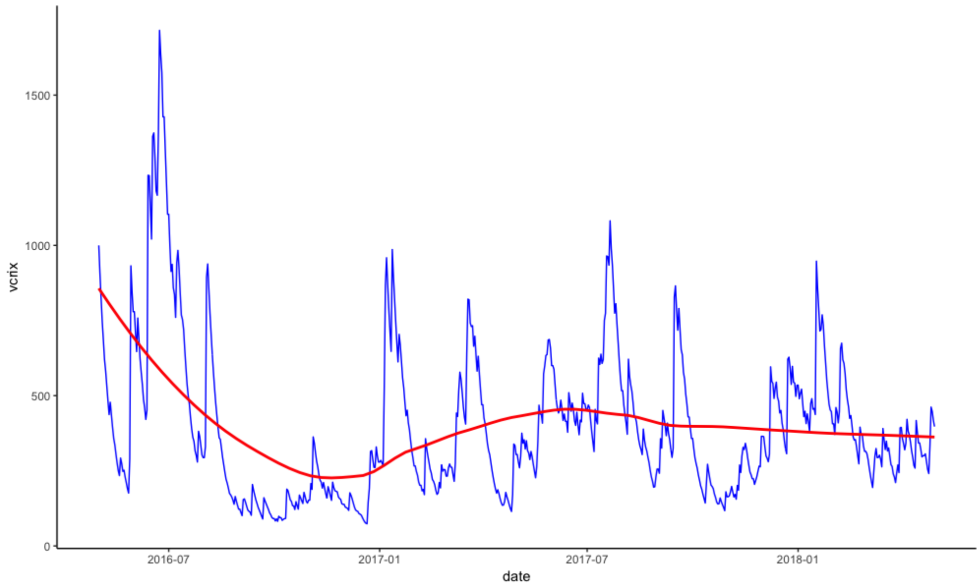

Figure 2.3: Trend in VCRIX.

VCRIXgraph

VCRIX captures the volatility jumps that corresponded to developments of crypto-space like spikes of interest in Bitcoin in the summer of 2016 (see more details in Figure 2.5), beginning of the first massive growth spike (also captured well by the CRIX in Figure 2.4), development of altcoins, SegWit fork in summer of 2017, as well as changes in legislative landscape like the Chinese ICO bans and stringent regulations in Korea.

Figure 2.5 also shows decreasing dependency of the index from the BTC returns as the latter was losing its dominating positions on the market (40% of the total market volume by the end of May 2018 CoinMarketCap (2018)). This statement is validated by Figure 2.6 that displays the change of BTC weight in CRIX over time.

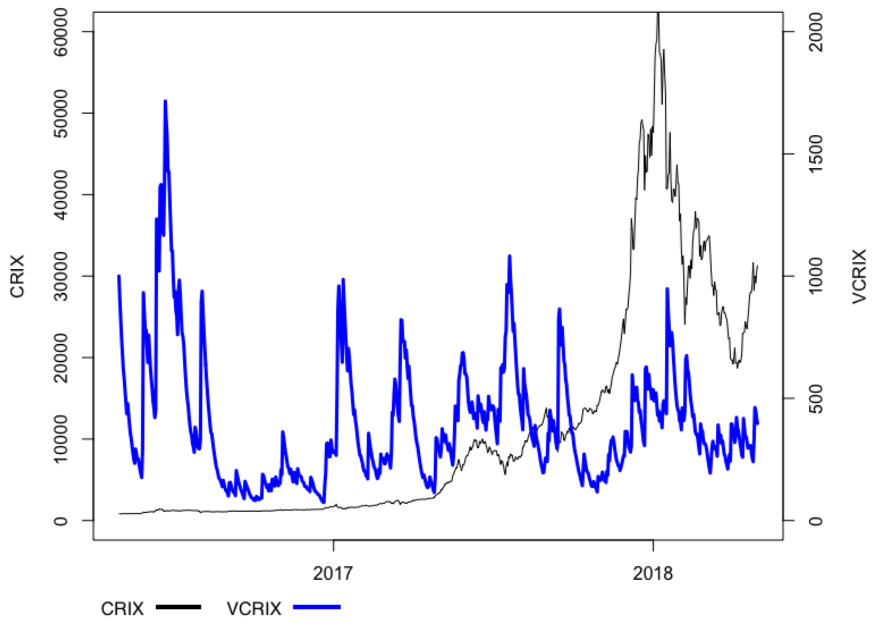

Figure 2.4: CRIX vs VCRIX.

VCRIXgraph

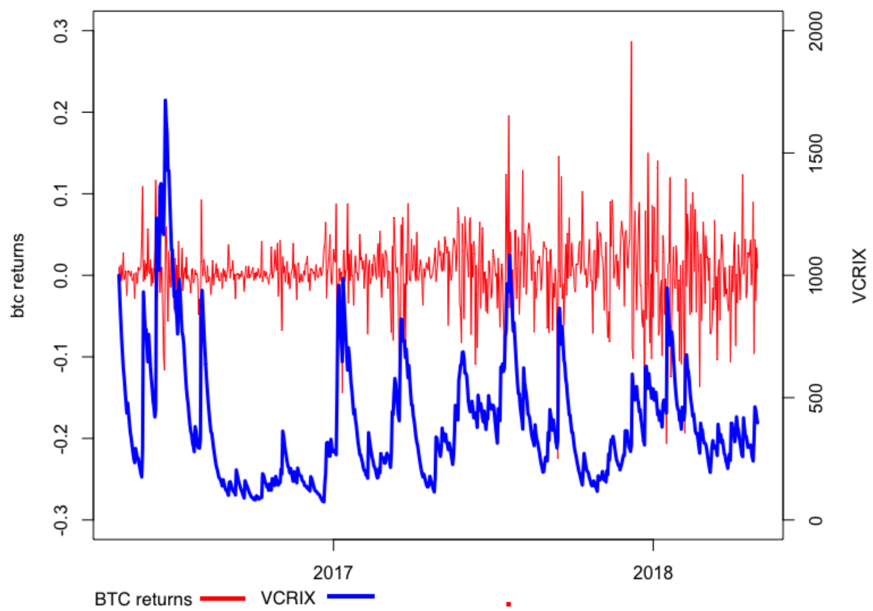

Figure 2.5: BTC returns vs VCRIX.

VCRIXgraph

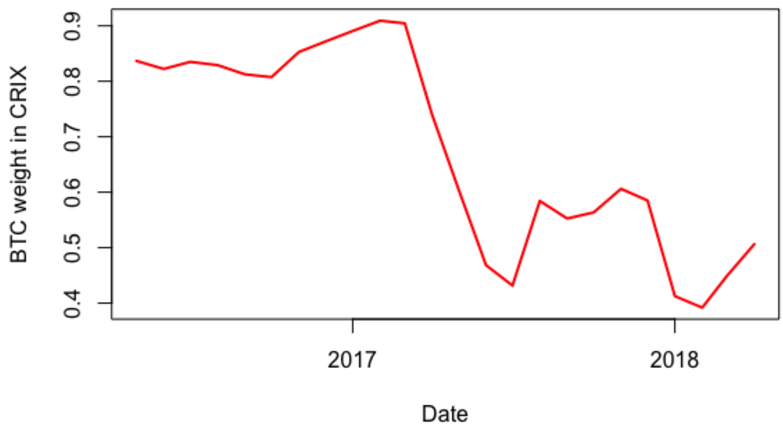

Figure 2.6: Change of BTC weight in CRIX.

2.5 Notes on lambda

By the end of the 90s JPMorgan et al. (1996) offered a recommendation for using decay factors of 0.94 (for trading, 74 day cutoff 1%) and 0.97 (for investing, 151 day cutoff at 1%). However, even lower values of λ might be applicable to the crypto markets due to its nature of extreme volatilities, not bounded by any regulation. The EWMA method dictates that lambda may be any value from 0 to 1. This paper will explore potential implications of selecting different values of λ.

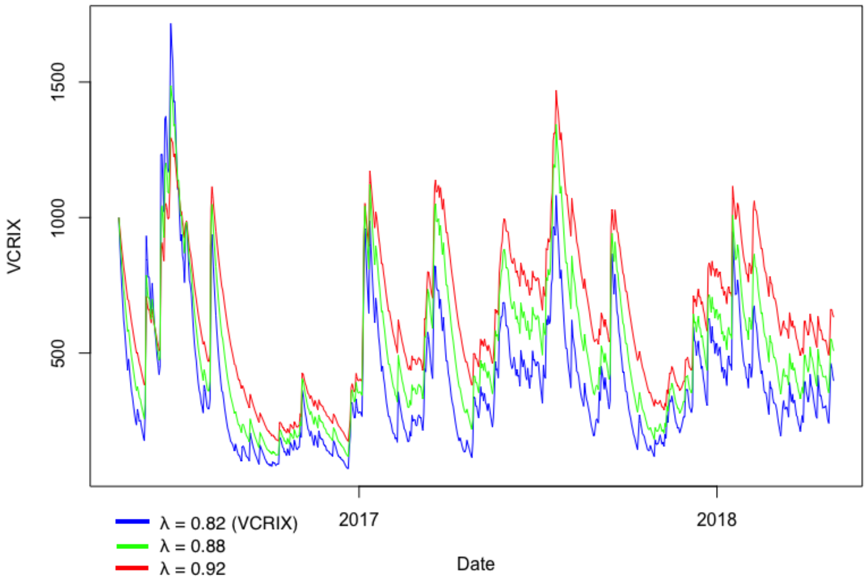

Figure 2.7: Comparison of VCRIX with different lambda values.

Figure 2.7 showcases the look of VCRIX with different decay parameters. As one can see, lower lambda is associated with less volatility clustering and faster "cooling" of the market aftershocks. While higher lambda allows for longer "agitation" periods, maintaining the high levels of volatility. λ of 0.86 was selected for VCRIX estimation, the rational will be explained further in subsection 3.1, however, I acknowledge the limitation of the fixed lambda approach. Short-term volatility is more responsive to shifts in market condition, long-term volatility offers more stability The further research will make an attempt to estimate the adaptive value of λ parameter according to look back length that becomes a function of market conditions. Switching between short or longer volatility can potentially achieve a better trade-off between responsiveness and smoothness which can lead to better outcomes when transaction costs become an issue (Härdle et al. (2003)).

3 Performance analysis

In order to evaluate the index and the chosen decay parameter two benchmarks were applied: estimation of true volatility with high-frequency intra-day CRIX data using Two Scale Realized Volatility (TSRV) method, and simulation of VIX using the VCRIX methodology. Within the two benchmark studies 3 major metrics were used: correlation, mean directional accuracy (MDA), and granger causality test in case of VIX simulation. These metrics were chosen with regard to the main purpose of the newly created index, which is grasping the fluctuation of the market sufficiently accurate in comparison to existing volatility estimates (high-frequency volatility). The MDA metrics is often used by economists in order to address differences in directional movement of the variable of interest.

$$\frac{1}{N}\sum_{t}\mathbf{1}_{sign(r_t - r_{t-1}) = sign(f_t - f_{t-1})} \tag{3.1}$$

where r_t is the actual value at time t and f_t is the forecast/estimated value at time t. Variable N represents the number of observations, $\mathbf{1}$ being an indicator function, which equals 1 when the argument in parenthesis is true and 0 otherwise, $sign()$ being a sign function.

$$\operatorname{sign}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases} \tag{3.2}$$

This method was tested on currency trading by Moosa and Vaz (2015) who advocate the superior importance of movement direction over magnitude. Like the correlation measurement, it will be applied to both high-frequency benchmark and the VIX simulation.

The Granger (1980) causality test was proposed as a statistical hypothesis test for determining whether one-time series is useful in forecasting another. A time series X is considered to Granger-cause time series Y if it can be shown, that those X values provide statistically significant information about future values of Y (defined through t-tests and F-tests on lagged values of X).

Within the scope of this paper the forecasting abilities of the VCRIX will be left aside.

3.1 High-frequency benchmark

The simplest approach to estimating volatility is to use standard deviation of historical returns. The backward-looking method has certain limitations, in particular, methods of risk evaluation built on backward-looking methods are always biased due to the usual violation of normality assumption according to Koopman et al. (2005). Volatility constructed from high-frequency observations could be used for the estimation of true volatility. This method was applied by Ait-Sahalia and Yu (2008) and proved sufficiently successful when dealing with currency exchange pairs. The Two Scale Realized Volatility (TSRV) method for handling the micro-structural noise (that would be expected in high-frequency estimation) proposed by Zhang et al. (2005) and developed further by Zu and Boswijk (2014), was considered an appropriate benchmark. The TSRV estimator is based on sub-sampling, averaging and bias-correction. It partitions the sample of size n into K subsamples. In this case$K$=40, T being number of time periods. This way the moving window will be $t_{41} - t_1$, $t_{42} - t_2$, estimating the realized volatility over the values of the window:

$$[r, r]_T^K = \frac{1}{K} \sum_{i=1}^{n-K+1} (r_{t_{i+K}} - r_{t_i})^2 \tag{3.3}$$

$$TSRV = (1 - \frac{z}{n})^{-1} [r, r]_T^K - \frac{z}{n} [r, r]_T^{all} \tag{3.4}$$

$$z = \frac{(n - K + 1)}{K} \tag{3.5}$$

In all three equations r is log returns on cryptos, while the first component of equation 3.4 is a coefficient to adjust for finite sample bias. As expected, there is an almost linear relationship between the value of lambda and the correlation of VCRIX with high-frequency benchmark (Table 3.1). Some gaps in the data can be seen in Figure 3.1 due to the down server times of CRIX. The Mean Directional Accuracy is the highest at λ=0.82, amounting to 0.46.

Value of lambda	Correlation	MDA
0.90	0.377	0.452
0.88	0.388	0.446
0.86	0.399	0.448
0.84	0.408	0.454
0.82	**0.416**	**0.462**
0.80	0.424	0.461
0.78	0.431	0.452
0.76	0.437	0.454
0.74	0.443	0.455
0.72	0.448	0.441
0.70	0.452	0.443

Table 3.1: Correlation and MDA between high-frequency CRIX volatility estimation and VCRIX.

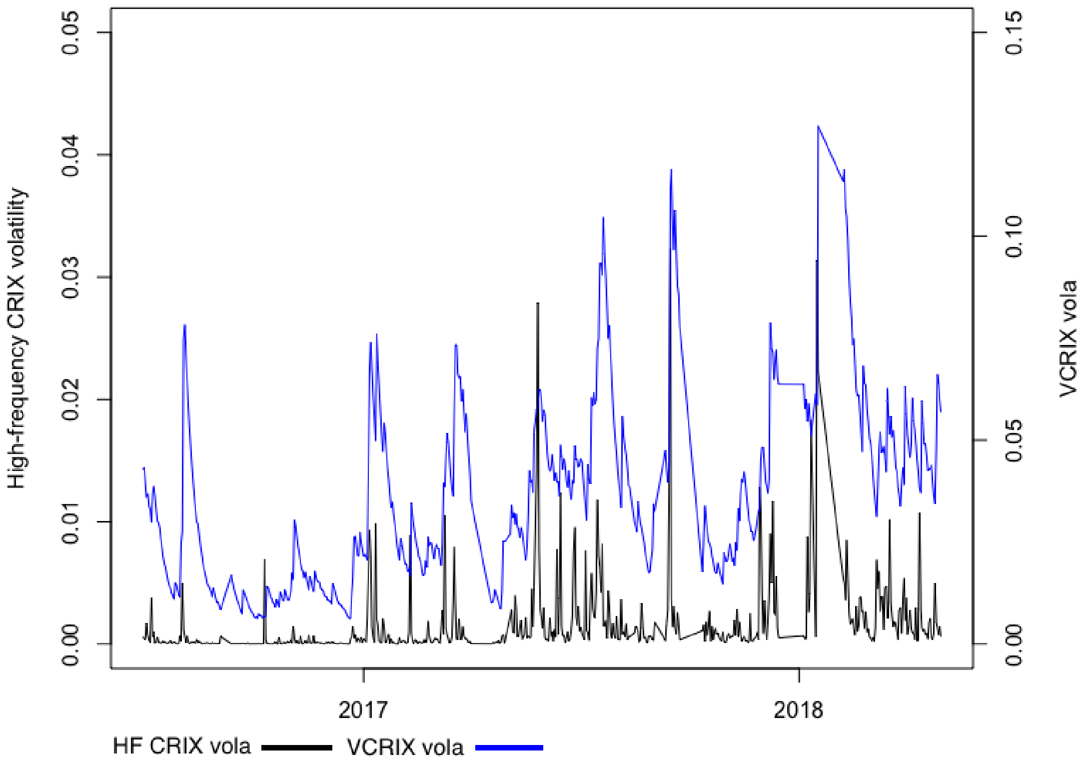

Figure 3.1: HF volatility estimation of CRIX vs VCRIX.

Additionally, a Johansen (1991) test was performed in order to analyze the time series for cointegration. The test showed a strong evidence for cointegration at 1% of significance level. At 10% significance level it can be concluded that there is stationary linear combination of HF benchmark and volatility underlying VCRIX, meaning that both times series are tracking the same underlying phenomena.

As can be seen from Figure 3.1, even with a relatively low value of decay parameters, VCRIX maintains a high level of estimated volatility much longer than the high-frequency benchmark. This feature can potentially yield the benefit of providing a higher level of alertness in case of the spike, as in most cases the spike would appear more than once in a short period of time.

3.2 VIX simulation

As mentioned in section 1.3 VCRIX models the volatility of the market and grasp the investor expectations as close as possible given the absence of the developed derivative market for the crypto-currencies. Its closest counterpart in traditional finance would be VIX. As it can be observed in Figure 3.2, the implied volatility is showing a different behavioral pattern than monthly standard deviation, not only time-wise (which is expected given that implied volatility is forward-looking) but also in the way it treats the events. the blue graph is exhibiting less brusque drops in volatility while the VIX drops down relatively fast after a shock. A similar behavior would be expected from a VCRIX and, as it was shown in subsection 3.1 it can achieved partially through lower decay parameters.

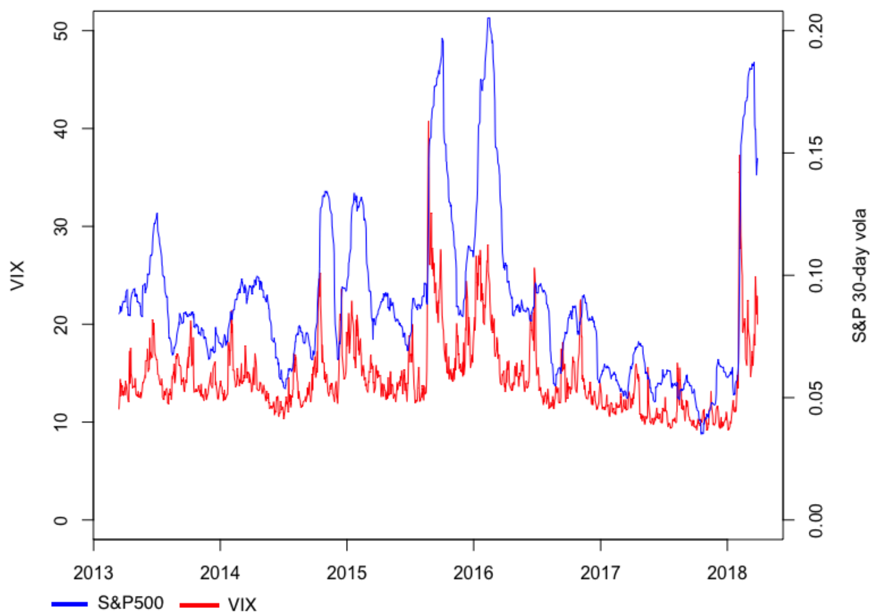

Figure 3.2: VIX and 30-day S&P 500 volatility Garman and Klass (1980).

In the conducted simulation the daily returns of 500 components of S&P 500 from September 2016 to February 2018 (359 observations) were sourced from YahooFinance (2018) and supplied into the VCRIX formula replacing the returns on crypto-currencies. The weights were calculated using the changing market cap of the companies and overall index capitalization. At this point, the *Divisor* was excluded from the formula yielding an estimation of the historical volatility. Additional complexity was introduced by the choice of the decay parameter λ. Figure 3.4 shows the comparative plot of the two time-series.

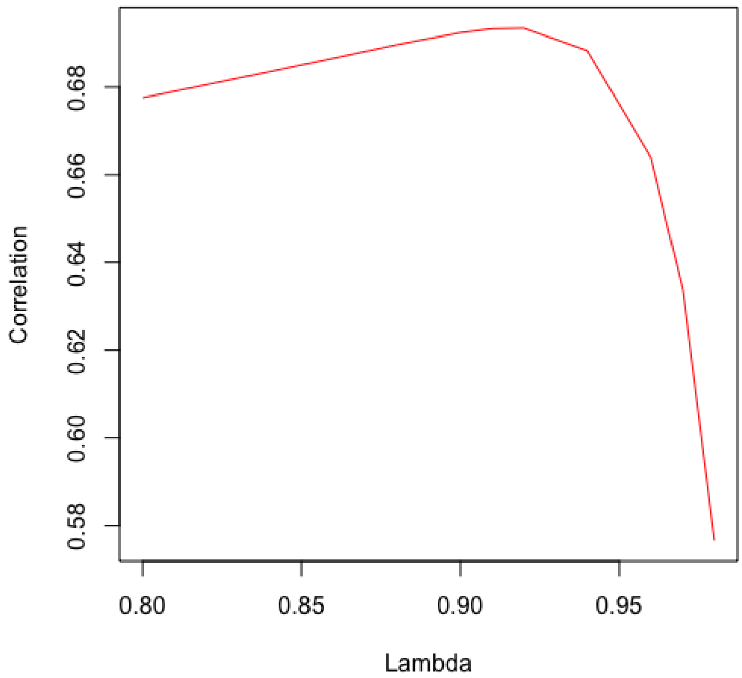

Figure 3.3: Correlation of VIX and estimated VIX with different lambda.

Given the λ=0.97, the original VIX and the simulated one have a correlation of 0.63. However, lower λ would provide higher correlations as seen in Figure 3.3, reaching its local peak at λ=0.92 with correlation 0.69.

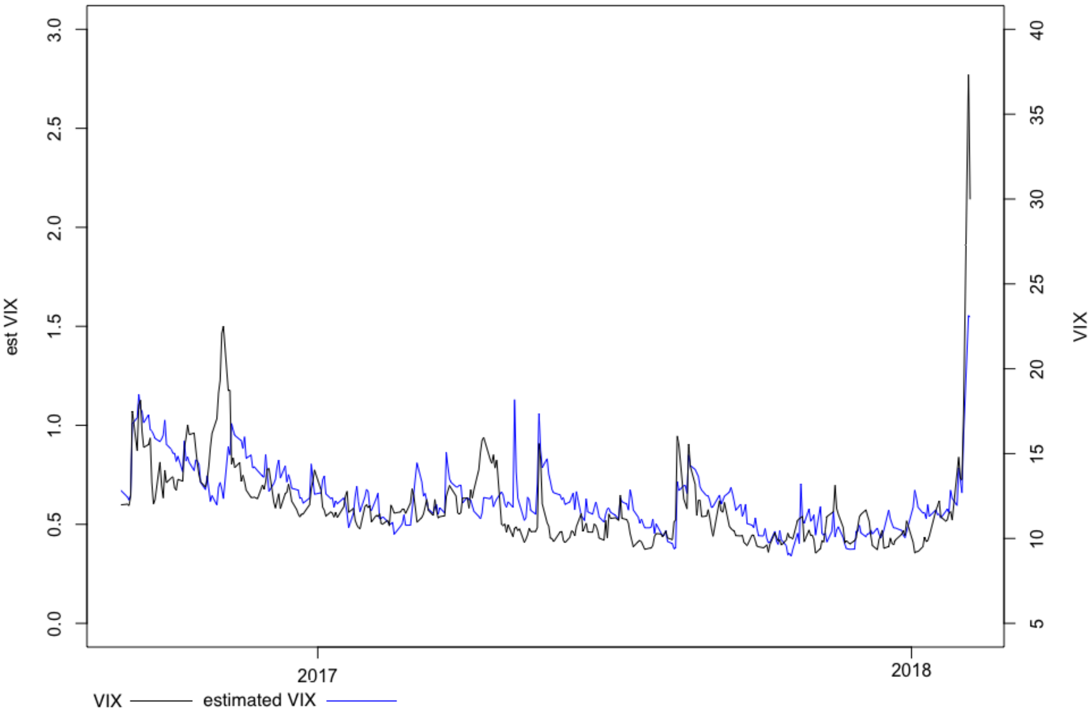

Figure 3.4: VIX and VIX simulation using VCRIX methodology.

It should be noted that natural lag is occurring as VIX is including options with maturities of 16-44 days. Thus a Granger causality test was applied showing the following results: at 1% significance level (p value =0.0002863) estimated volatility is caused by VIX with a 22 period lag (according to the usual maturity period for the S&P 500 swaps). Another important measurement shows that VIX and estimated VIX displayed MDA of 0.6 which captures the ability of VCRIX methodology to model the dynamics of the market close to VIX.

According to performed tests, VIX could be replicated using the VCRIX methodology and preserve the information about the market dynamics, however, further adjustments would be required for robust prediction.

4 Conclusion

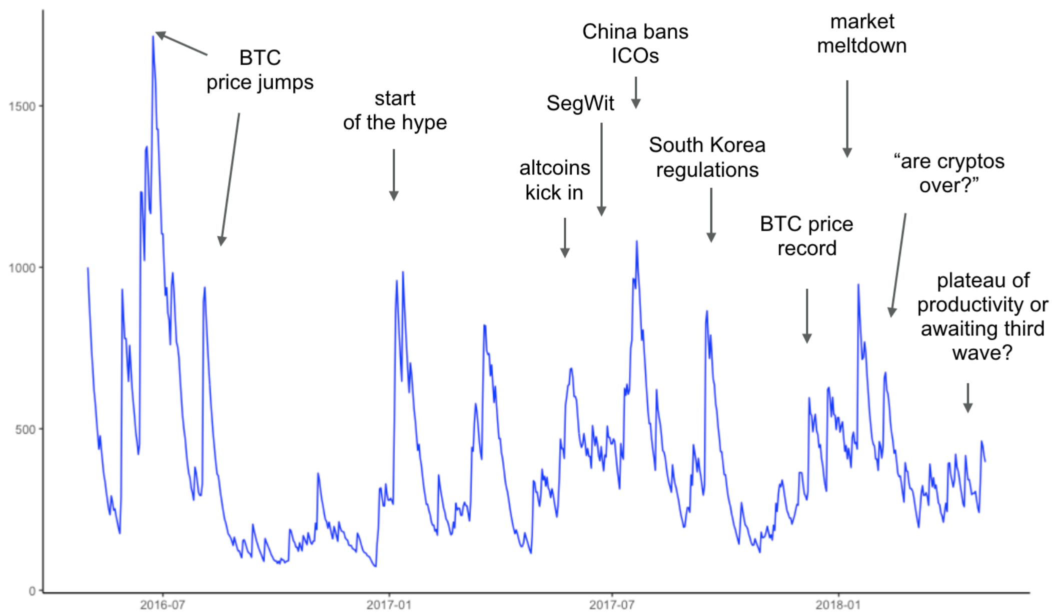

Figure 4.1: VCRIX interpretation.

The development of crypto-currencies happens at an unprecedented pace. They managed to become a new asset class taking a stand next to gold and stocks, gradually conquering new heights like the derivative market. CRIX index developed in 2016 became one of the first successful attempts to capture and communicate the state of the new market. VCRIX is an attempt to take this effort to the next level and offer operational tools for the integration of crypto-currencies into the established financial structure. VCRIX offers an estimate of the implied volatility in absence of the developed derivative market, bridging the gap for the implementation of the option pricing techniques. EWMA method used for the estimation of the variance-covariance matrix of the index components allowed to capture the relationship between the returns on cryptos and account for the integration effects. Additionally, a different approach to the decay parameter selection was offered and tested. In order to evaluate the proposed method, VIX was replicated using the components of S&P and VCRIX methodology. The estimated index showed a significant correlation with the actual VIX, granger causality and a substantial MDA of 0.6. Further development of the VCRIX method will be undertaken in order to improve the descriptive power, namely adaptive λ parameter for variance estimation, as well as introduction of skewed EWMA technique to account for the skewness in returns distribution. Additionally, predictive capabilities of VCRIX will be further tested. Given the financial theory, one has to expect the third large wave of volatility which nonetheless remains hard to predict using conventional methods. All in all, VCRIX has proven to be a valid method for the capturing of the CRIX behavior, superior to the straightforward estimation of historical volatility.

Bibliography

Ait-Sahalia, Y. and J. Yu (2008). High frequency market microstructure noise estimates and liquidity measures.

Andersen, T. G. and O. Bondarenko (2007, September). Construction and interpretation of model-free implied volatility. Working Paper 13449, National Bureau of Economic Research.

Baek, C. and M. Elbeck (2015). Bitcoins as an investment or speculative vehicle? a first look. *Applied Economics Letters 22*(1), 30–34.

Biktimirov, E. N. and C. Wang (2017). Model-based versus model-free implied volatility: Evidence from north american, european, and asian index option markets. *The Journal of Derivatives 24*(3), 42–68.

Black, F. and M. Scholes (1976). Taxes and the pricing of options. *The Journal of Finance 31*(2), 319–332.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Britten-Jones, M. and A. Neuberger (2000). Option prices, implied price processes, and stochastic volatility. *The Journal of Finance 55*(2), 839–866.

Catania, L., S. Grassi, and F. Ravazzolo (2018). Predicting the volatility of cryptocurrency time–series.

CBOE (2009). The cboe volatility index-vix. *White Paper*, 1–23.

Chen, C. Y., W. K. Härdle, A. J. Hou, and W. Wang (2018). Pricing cryptocurrency options: the case of crix and bitcoin.

Chen, S., C. Y.-H. Chen, W. Härdle, T. Lee, and B. Ong (2016). A first econometric analysis of the crix family. SFB 649 Discussion Paper 2016-031.

Cheung, A., E. Roca, and J.-J. Su (2015). Crypto-currency bubbles: an application of the phillips–shi–yu (2013) methodology on mt. gox bitcoin prices. *Applied Economics 47*(23), 2348–2358.

CoinMarketCap (2018). Charts. https://coinmarketcap.com/charts/. [Online; accessed 01-April-2018].

Demeterfi, K., E. Derman, M. Kamal, and J. Zou (1999). A guide to volatility and variance swaps. *The Journal of Derivatives 6*(4), 9–32.

Derman, E. and I. Kani (1994). Riding on a smile. *Risk 7*(2), 32–39.

Dimson, E. and P. Marsh (1990). Volatility forecasting without data-snooping. *Journal of Banking & Finance 14*(2-3), 399–421.

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*(6), 1343–1376.

Elendner, H., S. Trimborn, B. Ong, and T. M. Lee (2016). The cross-section of crypto-currencies as financial assets: an overview. Technical report, SFB 649 Discussion Paper.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Fengler, M. R., W. K. Härdle, and C. Villa (2003, Oct). The dynamics of implied volatilities: A common principal components approach. *Review of Derivatives Research 6*(3), 179–202.

Garman, M. B. and M. J. Klass (1980). On the estimation of security price volatilities from historical data. *Journal of business*, 67–78.

Glaser, F., K. Zimmermann, M. Haferkorn, M. Weber, and M. Siering (2014). Bitcoin-asset or currency? revealing users' hidden intentions.

Granger, C. W. (1980). Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and control 2*, 329–352.

Härdle, W., H. Herwartz, and V. Spokoiny (2003). Time inhomogeneous multiple volatility modeling. *Journal of Financial econometrics 1*(1), 55–95.

Hayes, A. S. (2017). Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telematics and Informatics 34*(7), 1308–1321.

Hull, J. and A. White (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of risk 1*(1), 5–19.

Hyndman, R. J. and H. L. Shang (2009). Forecasting functional time series. *Journal of the Korean Statistical Society 38*(3), 199–211.

icowatchlist (2018). Charts. icowatchlist.com. [Online; accessed 01-April-2018].

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society*, 1551–1580.

JPMorgan et al. (1996). Riskmetrics technical document.

Koopman, S. J., B. Jungbacker, and E. Hol (2005). Forecasting daily variability of the s&p 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance 12*(3), 445–475.

Kuen, T. Y. and T. S. Hoong (1992, Apr). Forecasting volatility in the singapore stock market. *Asia Pacific Journal of Management 9*(1), 1–13.

Lu, Z., H. Huang, and R. Gerlach (2010). Estimating value at risk: from jp morgan's standard-ewma to skewed-ewma forecasting. *University of Sydney*.

Madan, D. B., P. P. Carr, and E. C. Chang (1998). The variance gamma process and option pricing. *Review of Finance 2*(1), 79–105.

Markowitz, H. (1952). Portfolio selection. *The journal of finance 7*(1), 77–91.

Moosa, I. and J. Vaz (2015). Directional accuracy, forecasting error and the profitability of currency trading: model-based evidence. *Applied Economics 47*(57), 6191–6199.

Neubürger, H.-J. (1994). Einsatz (derivativer) finanzinstrumente in der praxis. *Institut der Wirtschaftsprüfer, Bericht über die Fachtagung*, 311–340.

Siriopoulos, C. and A. Fassas (2009). Implied volatility indices–a review.

Trimborn, S. and W. K. Härdle (2016). CRIX an Index for blockchain based Currencies. *CRC 649 Discussion Paper 2016-021, revise and resubmit Journal of Empirical Finance*.

Tse, Y. K. (1991). Stock returns volatility in the tokyo stock exchange. *Japan and the World Economy 3*(3), 285–298.

White, L. H. (2015). The market for cryptocurrencies. *Cato J. 35*, 383.

YahooFinance (2018). Financial Data. https://finance.yahoo.com. [Online; accessed 15-April-2018].

Yermack, D. (2015). Is bitcoin a real currency? an economic appraisal. pp. 31–43.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

Zu, Y. and H. P. Boswijk (2014). Estimating spot volatility with high-frequency financial data. *Journal of Econometrics 181*(2), 117–135.

Declaration of Authorship

I hereby confirm that I, Alisa Kolesnikova, have authored this master thesis independently and without use of others than the indicated sources. Where I have consulted the published work of others, in any form (e.g. ideas, equations, figures, text, tables), this is always explicitly attributed.

Berlin, May 30, 2018

Alisa Kolesnikova

Hiermit erklare ich, Alisa Kolesnikova, dass ich die vorliegende Arbeit allein und nur unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe. Die Prüfungsordnung ist mir bekannt. Ich habe in meinem Studienfach bisher keine Masterarbeit eingereicht bzw. diese nicht endgültig nicht bestanden.

Berlin, May 30, 2018

Alisa Kolesnikova